UNITED STATES PATENT APPLICATION

FOR

A METHOD AND APPARATUS FOR PROVIDING MULTIPLE LEVELS OF ABSTRACTIONS IN DESCRIPTIONS OF AUDIOVISUAL CONTENT

INVENTORS:

HAWLEY K. RISING III

PREPARED BY:

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN, LLP
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1026

(408) 720-8598

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number _____ EL431888471US _____

Date of Deposit _____ July 13, 2001 _____

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

_____ MICHELLE BEGAY _____

(Typed or printed name of person mailing paper or fee)

_____ Michelle Begay _____

(Signature of person mailing paper or fee)

A METHOD AND APPARATUS FOR PROVIDING MULTIPLE LEVELS OF ABSTRACTION IN DESCRIPTIONS OF AUDIOVISUAL CONTENT

RELATED APPLICATIONS

5 This application claims the benefit of U.S. Provisional Application No. 60/219,164, filed July 19, 2000.

FIELD OF THE INVENTION

The present invention relates to multimedia data processing and more specifically to processing of multimedia data based upon semantic descriptions.

BACKGROUND OF THE INVENTION

10 The Motion Picture Expert Group (MPEG) develops standards concerning audiovisual content. One component of the MPEG standard scheme includes MPEG-7 standards which are directed to providing descriptions of audiovisual content that may be of interest to the user. Specifically, the MPEG-7 standards are

15 developed to standardize information describing the audiovisual content. The MPEG-7 standards may be used in various areas, including storage and retrieval of audiovisual items from databases, broadcast media selection, tele-shopping, multimedia presentations, personalized news service on the Internet, etc.

According to current MPEG-7 standards, descriptions of audiovisual

20 content are divided into structural (or segment) descriptions and semantic descriptions. Structural descriptions describe the audiovisual content from the viewpoint of its structure. That is, the descriptions are structured around segments which represent physical spatial, temporal or spatio-temporal components of the audiovisual content. Each segment may be described by

signal-based features (color, texture, shape, motion, audio features, etc.) and some elementary semantic information.

Semantic descriptions describe the audiovisual content from the viewpoints of its conceptual notions. The semantic descriptions involve entities

5　such as objects, events, abstract concepts and relationships. The structural descriptions and semantic descriptions are related by a set of links, which allows the audiovisual content to be described on the basis of both content structure and semantics together. The links relate different semantic concepts to the instances within the audiovisual content described by the structural descriptions.

10　Semantic descriptions describe entities that are either concrete or abstract. A concrete entity can have an instance in the real world or the media. An abstract entity results from applying abstraction to a physical entity. Various levels of abstraction may be applied to a physical entity when creating a description. For instance, a description may be created as a media abstraction. A media

15　abstraction results from generalizing a media, i.e., the description is not attached to a specific media. A formal abstraction (or a lambda abstraction) is created from a description of a concrete entity by generalizing one or more elements of the entity, i.e., by substituting one or more elements of the entity by a variable. A description may also be created as an abstraction of a higher degree (e.g., a

20　metaphor, an abstraction of abstractions, etc.).

An abstraction can be used as a template for creating descriptions of specific audiovisual items or lower level abstractions. For instance, an archivist who needs to create a catalog of a large number of audiovisual items contained in a digital library can use such templates to create descriptions of these audiovisual

080398.P426　　　　　　　　　-2-

items. In another example, existing abstractions can be reused by software when generating new descriptions. However, no mechanism currently exists to indicate that a description is an abstraction and to identify the type of the abstraction (i.e., whether the description is a media abstraction, a lambda abstraction, a metaphor, etc.), thereby complicating a task of creating new descriptions from existing descriptions.

SUMMARY OF THE INVENTION

A method and apparatus for processing descriptions of audiovisual content are described. According to one embodiment, a description of audiovisual content is evaluated and a determination is made as to whether the description is an abstraction. If it is determined that the description is an abstraction, then a level of abstraction is calculated and an indicator of the level of abstraction is stored with the description of audiovisual content.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation in the figures of the accompanying drawings in which like reference numerals refer to similar elements.

5   **Figure 1** is a block diagram of one embodiment of a system for processing descriptions of audiovisual content;

Figure 2 is a flow diagram of one embodiment for processing descriptions of audiovisual content;

Figure 3 illustrates two usage modes of semantic descriptions; and

10   **Figure 4** is a block diagram of one embodiment of a computer system.

DETAILED DESCRIPTION

A method and apparatus for processing descriptions of audiovisual content are described. In the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough

5   understanding of the present invention. However, it will be apparent to one skilled in the art that the present invention may be practiced without these specific details. In some instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

10   Some portions of the detailed descriptions that follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is

15   here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times,

20   principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise

as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that

5   manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

The present invention also relates to apparatus for performing the

10   operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs,

15   and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

The algorithms and displays presented herein are not inherently related to

20   any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described

with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

Figure 1 is a block diagram of one embodiment of a system 100 for processing descriptions of audiovisual content. System 100 includes a data repository 120 coupled to audiovisual content 110. Data repository 120 contains descriptions of audiovisual content 110. In one embodiment, audiovisual content 110 is stored independently. Alternatively, audiovisual content may be stored together with the descriptions in data repository 120.

The descriptions consist of descriptors and description schemes. Descriptors represent features of the audiovisual content 110 and define the syntax and the semantics of each feature representation. Description schemes (DS) specify the structure and semantics of the relationships between their components. These components may be both descriptors and description schemes. Conceptual aspects of a description scheme can be organized in a tree or in a graph. The graph structure is defined by a set of nodes that represent elements of a description scheme and a set of edges that specify the relationship between the nodes.

System 100 further includes a description processor 130. In one embodiment, the description processor 130 evaluates descriptions stored in data repository 120 and determines which of these descriptions are abstractions. A description is an abstraction if it is not linked to a specific entity or a specific media. If a description is an abstraction, the description processor 130 determines its abstraction level. An abstraction level identifies a type of abstraction. For

instance, a description may be a media abstraction, which is created by generalizing the media. Alternatively, a description may be a formal abstraction (or a lambda abstraction) which generalizes one or more elements of a concrete entity by replacing them with variables. A description may also be an abstraction

5   of a higher degree (e.g., a metaphor, an abstraction of graph relations, etc.). When the abstraction level is determined, it is stored in data repository 140, together with the description.

    In another embodiment, the description processor 130 makes a determination of whether a description is an abstraction during creation of the

10   description. If the determination is positive, the description processor 130 calculates the level of abstraction, and the new description is stored in data repository 140 with an indicator of the level of abstraction. In one embodiment, a new description is created using an existing description. For example, software that generates descriptions can use an existing description of an abstract event to

15   create a new description of a concrete event by replacing one or more variables in the existing description with corresponding attributes of the concrete event. In this embodiment, an abstraction level indicator may be used to find an existing description that can be used to create a new description.

    System 100 also includes a search module 140 which is responsible for

20   searching data repository 120 and retrieving requested descriptions and audiovisual content 110. In one embodiment, the search module 140 retrieves descriptions from data repository 120 based on a user selection of an abstraction level.

Figure 2 is a flow diagram of one embodiment for processing descriptions of audiovisual content. At process block 204 a description of audiovisual content is evaluated. In one embodiment, the description being evaluated is an existing description which is stored in data repository 120. Alternatively, the evaluation is

5   performed when creating a new description.

At processing block 206, a determination is made as to whether the description is an abstraction. This determination depends on whether the description is linked to a specific media or a concrete audiovisual item.

If the determination is positive, a level of abstraction is calculated at

10   processing block 208. For example, the lowest level of abstraction may correspond to a media abstraction, the next level may correspond to a lambda abstraction, and higher levels of abstraction may correspond to metaphors or abstractions of abstractions. A metaphor describes a person or object in a literary way by referring to something that is considered to possess similar characteristics

15   to the person or object being described. For example, such expressions as "flying as a bird" and "feeling like a fish out of water" are metaphors. Abstractions of abstractions include abstractions of graph relations, graph morphisms, etc. Abstractions of graph relations are made by describing sets of relations, or by using a very general relation in place of something more specific. Abstractions of

20   graph morphisms are essentially descriptions of how the descriptions themselves relate to each other. Further, the description may be stored in data repository 120, together with an indicator of the level of abstraction.

In one embodiment, an abstraction level indicator is an optional element of a description, i.e., it is added to the description only if the description is an abstraction. In this embodiment, if the determination made at decision box 206 is negative (i.e., the description is not an abstraction), the description is stored in

5   data repository 120 without an abstraction level indicator at processing block 212. In another embodiment, an abstraction level indicator may be a required element of a description that includes a value identifying an abstraction level of the description, or alternatively, indicating that the description is not an abstraction.

In one embodiment, the abstraction level indicator is utilized to search data

10   repository 120 and retrieve descriptions of a certain level of abstraction. In another embodiment, the abstraction level indicator is utilized to create new descriptions from an existing description as described in more detail above.

In one embodiment, descriptions of audiovisual content include structural descriptions and semantic descriptions. Structural descriptions describe the

15   audiovisual content from the viewpoint of its structure, i.e., such components of the audiovisual data as regions, segments, frames, shots, etc. The structural descriptions may describe such perceptual features of components of the audiovisual data as color, texture, shape, motion, pitch, rhythm, etc. Semantic descriptions describe the audiovisual content from the conceptual viewpoints, i.e.,

20   they describe the meaning of audiovisual content. The semantic descriptions involve entities such as objects, events, states and abstract concepts. The structural descriptions and semantic descriptions are related by a set of links, that allows the audiovisual content to be described on the basis of both content structure and semantics together. The links relate different semantic concepts to

the instances within the audiovisual content described by the structural descriptions.

According to the MPEG standards, the semantic descriptions are divided into two types depending on their usage mode. The first type includes data

5 describing audiovisual content and is referred to as "descriptive data". For instance, the descriptive data may include a set of descriptions (i.e., descriptors and description schemes) describing a specific instance in the audiovisual material. The second type is used to access one or more individual instances in the audiovisual content and is referred to as "accessing data". For instance, the

10 object "car" may be used to locate all instances of cars in the audiovisual content. That is, accessing data functions as an index, pointing to all occurrences of a particular description. In one embodiment, each semantic description includes an element referred to as a "usage label", which indicates whether the description is of an accessing type or a descriptive type. In one embodiment, the usage label is

15 used for abstraction. Alternatively, any other element (existing or new) of a description may be used for this purpose.

Figure 3 illustrates two usage modes of semantic descriptions. Referring to **Figure 3**, descriptive data 310 includes a set of descriptions D1 through D3, wherein each of the description D1 through D3 describes a specific piece of

20 audiovisual content 330 using descriptions of associated attributes. For example, audiovisual content 330 may be a video of a hockey game. Then, D1 may describe a first period of the hockey game using descriptions d11 through d1n, D2 may describe a second period of the hockey game using descriptions d21 through d2n, and D3 may describe a third period of the hockey game using descriptions d31

through d3n.

Accessing data 320 includes a set of descriptions a1 through an, wherein each accessing data 320 is linked to all occurrences of any of descriptive data 310 in the audiovisual content 330. For example, description a1 may point to all occurrences of scoring a goal in the soccer game. Scoring of the goal may be described by descriptions d11, d23 and d3n.

However, if the arrows pointing from an accessing description to instances in the audiovisual content 330 are reversed, the accessing description becomes an abstraction of a set of descriptions within the descriptive data 310. That is, each descriptive data pointing to an accessing description would have an "is a" or a similar relation to this accessing description. For instance, if the arrows pointing from a1 to instances of the goal in the audiovisual content are reversed, the accessing description a1 becomes an abstraction of a description set consisting of d11, d23 and d3n.

In one embodiment, a usage label is an optional element of a description. For example, if a description is of an accessing type, it is considered an abstraction. Therefore, accessing descriptions always include a usage label. If a description is of a descriptive type, it may or may not include a usage label depending on whether it refers to a specific media or a specific instance in the media (i.e., depending on whether the description is an abstraction or not).

In one embodiment, the usage label includes a numeric field and a "term". The numeric field identifies the abstraction type of the description. For instance, the value of 0 may correspond to a media abstraction, 1 may correspond to a lambda abstraction, 2 may correspond to a metaphor, 3 may correspond to an

abstraction of graph relations, 4 may correspond to an abstraction of graph morphisms, etc. The "term" is a pointer to a particular entry in a classification scheme. The classification scheme represents a controlling vocabulary which is used for creating descriptions. Each entry in the classification scheme includes a

5   set of related data (e.g., labels, definitions in different foreign languages, etc.). That is, the description is an abstraction of the set of data included in the classification scheme entry. Accordingly, one can learn additional information related to the description from the data included in the entry referred to by the term.

10   In another embodiment, any new or existing element rather than the usage label may be used to include a numeric field and a term in descriptions. In yet another embodiment, two elements are used for this purpose: a first element contains a numeric field and a second element contains a term.

Figure 4 is a diagrammatic representation of a machine in the exemplary

15   form of a computer system 400 within which a set of instructions, for causing the machine to perform any one of the methodologies discussed above, may be executed. In alternative embodiments, the machine may comprise a network router, a network switch, a network bridge, Personal Digital Assistant (PDA), a cellular telephone, a web appliance or any machine capable of executing a

20   sequence of instructions that specify actions to be taken by that machine.

The computer system 400 includes a processor 402, a main memory 404 and a static memory 406, which communicate with each other via a bus 408. The computer system 400 may further include a video display unit 410 (e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT)). The computer system 400

also includes an alpha-numeric input device 412 (e.g., a keyboard), a cursor control device 414 (e.g., a mouse), a disk drive unit 416, a signal generation device 420 (e.g., a speaker) and a network interface device 422.

The disk drive unit 416 includes a computer-readable medium 424 on

5 which is stored a set of instructions (i.e., software) 426 embodying any one, or all, of the methodologies described above. The software 426 is also shown to reside, completely or at least partially, within the main memory 404 and/or within the processor 402. The software 426 may further be transmitted or received via the network interface device 422. For the purposes of this specification, the term

10 "computer-readable medium" shall be taken to include any medium that is capable of storing or encoding a sequence of instructions for execution by the computer and that cause the computer to perform any one of the methodologies of the present invention. The term "computer-readable medium" shall accordingly be taken to included, but not be limited to, solid-state memories,

15 optical and magnetic disks, and carrier wave signals.

Thus, a method and apparatus for processing descriptions of audiovisual content have been described. Although the present invention has been described with reference to specific exemplary embodiments, it will be evident that various modifications and changes may be made to these embodiments without departing

20 from the broader spirit and scope of the invention. Accordingly, the specification and drawings are to be regarded in an illustrative rather than a restrictive sense.